Size Matters: The Effect of the Size of Ethnic Groups on Development

December 24, 2015

Abstract

Word Count: 12,423

The share of ethnic groups is one of the most important features of African politics. It affects civil wars, representation in government positions, distributive and allocative policies. In this paper we use the partition of ethnic groups as a natural experiment in order to estimate the effect of the share of these ethnic groups on development. We show that larger groups have an advantage in terms of development and that the partition in itself does not matter for development. This result is explained by the fact that the partition matters only when the resulting groups are relatively small, since their lack of political representation may weaken support for institutions, may bias policies and the provision of ethnic/regional public goods.

Keywords: Ethnicity; Partition; Political Representation; Regional Development

JEL Codes: O10

1 Introduction

In the last few years a large body of research has tried to explain why some countries still lag behind while other countries have enjoyed steady long-term economic development. Among the several explanations which have been offered, the institutional hypothesis has been the one which has dominated the economic arena. Poorly performing political and institutional structures, together with inefficient legal and court systems, are among the primary causes of poor development (Acemoglu, Johnson and Robinson, 2001, 2002, 2005; Glaeser, La Porta, de Silanes and Shleifer, 2004; La Porta, de Silanes, Shleifer and Vishny, 1997, 1998; etc.). With regard to political institutions, Acemoglu and Robinson (2012) distinguish between extractive and inclusive institutions. Unlike inclusive institutions, which are conducive to long-term development, extractive institutions are designed to serve the interest of small elites by exploiting the rest of the population and for this reason they are highly detrimental to development.

Africa is one of the continents which have suffered most in terms of the persistence of extractive institutions. Inefficient property rights, patronage politics, corruption, mistrust and unstable democratic institutions have long been proposed as sources of poor development in the continent. Some of these institutions are the result of the slave trade (Nunn 2008; 2010). Others derive from the legacy of colonization and the subsequent Scramble for Africa which at times has been considered even more harmful than colonization itself (Asiwaju 1985; Dowden, 2008).

Loosely speaking, the Scramble for Africa consists in the arbitrary and improper border design which partitioned a signicant fraction of the population belonging to existing ethnic groups. As a result, where political borders do not always coincide with pre-existing tribal institutions (Englebert, 2000a), large shares of the population belonging to different ethnic groups have been forced to coexist in artificial states (Alesina, Easterly, and Matuszesky, 2010). The discontinuity between pre- and post-colonial institutions has caused illegitimacy (Englebert, 2000a), civil conflicts (Fearon and Laitin, 2003; Fearon, 2004), ethno-political mobilization (Posner 2004a, 2004b), and particular rather than collective policies (Miguel, 2004). A somewhat more sophisticated interpretation of the effect of the Scramble for Africa is provided by Posner (2004a). Focusing on the partition of the Chewa and Tumbukas peoples between Malawi and Zambia, he argues that *"the political, social, and cultural salience of the cleavage depends on the sizes of the group that the cleavage defines relative to the sizes of the political and social arenas in which the groups are located"* (Posner, 2004a, p. 543). Thus, the idea that the political salience of a cultural cleavage results from the arbitrary imposition of boundaries in itself seems open to doubt.

In this paper we build on the hypothesis in Posner (2004a) in order to investigate the idea that the partitioning of ethnic groups in Africa matters only when the share of groups resulting from the imposition of national boundaries is small in relation to the political and social arenas in which the groups are located. To analyse the effect of partitioned groups (and their relative share) on development, we use data at group level from Murdock (1959, 1967) which we merge with ethno-country estimates on GDP from satellite imagery of the light density at night from the NOAA/NGDC (National Geophysical Data Center). We initially exploit the variation across groups to show that the negative effect of the partition is not robust to the inclusion of additional controls. After establishing that the effect of the partition is not significant, we exploit the variation within groups in order to estimate the effect of the share of the group on development. Exploiting the variation within groups allows us to control for confounding variables, given that partitioned groups belonging to the same historical homeland are likely to start from the same level of income per capita and, share the same culture and institutions, which then neutralizes biases of this kind. Using this sort of matching estimator we find a close relationship between the size of the partitioned group and development. The robustness of the results is tested using a series of different approaches. First, we use a regression discontinuity approach exploiting the discontinuity at the border. Then we re-estimate same models using an IV estimator in order to provide exogeneity to the share of partitioned groups and find that the results are quite close to those estimated using the regression discontinuity approach. Finally we change the dependent variable using household data from the Afrobarometer IV to check the sensitivity of the results to different measures of economic development and we still find a significant effect of the share of partitioned groups on proxies of individual income.

In order to shed light on some of the channels through which the share of ethnic groups may affect development, we test a set of different theories which have been proposed by the economic and political science literature. We evaluate the effect of the share of partitioned ethnic groups on informal institutions, investment in infrastructures, and trust (which we use to proxy community networks). Even though estimates are not causal we find support to the idea that smaller groups tend to provide more support than larger groups do to the informal institutions of ethnic leaders with regard to the allocation of land and responsibility for maintaining rule and order. This finding seems to be consistent with the idea of the persistence of customary institutions in rural areas which then may hamper development because of weak institutions and poor enforcement of property rights. This result is also supported by a significant effect of the share of groups on trust. In fact, human interactions become extremely important in absence of formal institutions. As a result groups which rely on informal institutions are likely to trust more too. We also find

a significant effect of the share of ethnic groups on investment in infrastructures which we proxy with the distance from railways and roads. Smaller groups are further than larger groups from railways and roads, making the former more likely to suffer from increased transportation costs and therefore retard regional development.

This paper is structured as follows. In the next Section we review part of the literature on economic development in Africa, focusing on papers related to the partition of ethnic groups there and its effect on development. In Section 3 we discuss the sources used to collect data and the way we have used these sources in order to carry out our analysis. In Section 4 we provide evidence in support of the idea that it is not the imposition of arbitrary boundaries that matters. At the same time we show a positive and significant effect of the shares of the partitioned groups on development. In Section 5 we conduct a series of robustness checks evaluating whether this effect is robust to alternative estimators (i.e. IV and RD approaches) and to changes in the dependent variables. In Section 6 we review and test some of the theories which may explain the relationship between the share of ethnic groups and development. The paper ends with a brief conclusion.

2 Related Literature

A significant part of the literature in political science has focused on ethnic politics and the impossibility of developing a nation-building process in a country when it has experienced ethnic cleavages (Horowitz, 1985; Huntington, 1996). The presence of ethnic cleavages leads to ethno-politic mobilization (Posner, 2004b) and hence politicians find it easier to build electoral support along ethnic lines (Eifart, Miguel and Posner, 2010). The result of ethno-politics is to foster ethno-culture and ethno-institutions, leading to a lack of confidence in national political institutions (Norris and Mattes, 2003).

This process seems to have a particularly severe effect on most countries in Africa (Mattes and Gouws, 1999; Mattes and Piombo, 2001; Norris and Mattes, 2003) where ethno-politics, the current ethnic diversity and the resulting weak institutions seem to be the result of the Scramble for Africa followed by the arbitrary imposition of state boundaries (e.g. Ajala, 1983; Asiwaju, 1985; Barbour, 1961; Bello, 1995; Brownlie, 1979; Davidson, 1992; Kum, 1993; Nugent and Asiwaju, 1996; Touval, 1966, Englebert, 2000a, 2000b).

Building on this idea, authors have used data on whether state boundaries are represented by a straight line and the length of these straight lines in order mainly to find a possible effect of artificial state boundaries on civil conflict (e.g. Clapham, 1996; Odugbemi,

1995; Ottaway, 1999; Touval, 1969; Bach, 1999; Nugent, 1996; Barbour, 1961; Bayart, 1996; Griffiths, 1996; Young, 1996; Herbst, 2000; Englebert, Tarango and Carter, 2002). However the evidence is somewhat mixed.

The effect of ethnic divisions has also been widely debated in the economic literature (Easterly and Levine, 1997; Alesina, Devleeschauwer, Easterly, Kurlat and Wacziarg, 2003; Fearon, 2003; Montalvo and Reynal-Querol, 2005; Esteban and Ray, 2008; Michalopoulos, 2012). Part of this literature has focused on the effect of pre-colonial and colonial ethnic institutions on development. Gennaioli and Rainer (2006) and Michalopoulos and Papaioannou (2013) document a close association between pre-colonial ethnic political centralization and measures of national and regional development. Michalopoulos and Papaioannou (2014a) show that differences in countrywide institutional structures across a national border do not explain within-ethnicity differences in economic performance, as captured by satellite images of light density. Michalopoulos and Papaioannou (2014b) find that support for traditional leaders in sub-Saharan Africa remains quite strong and the persistence of this *"dual"* institutional environment is associated with development divergences across areas. Heldring and Robinson (2012) compare differences in the administration of African colonies in order to evaluate the effect of colonization on development. Englebert (2000a) looks at the continuity between pre- and post-colonial institutions and finds in a cross-country growth analysis that institutional continuity explains the Africa dummy effect. To proxy artificial states, Alesina, Easterly and Matuszesky (2010) use measures for whether there are partitioned groups within the country and whether the country boundaries are represented by a straight line (which they proxy using bounding-boxes and a fractal measure).

With respect to this literature, the paper shows that unless a group is relatively small it is not the imposition of arbitrary boundaries that matters for development. Only in the latter case there is a negative effect on development and this effect is beyond any sort of effect of pre-colonial and current institutions for which we control. In addition, with respect to Alesina, Easterly and Matuszesky (2010), the paper focuses on ethnic groups rather than countries because it is normal to expect that the groups which have been partitioned are the ones which are most affected by the partition.

The paper is also closely related to the literature on ethnic groups, political representation, patronage politics and development. Rainer and Trebbi (2014) provide detailed data on ministerial positions and ethnic groups over a 40-year period in 15 democratic and autocratic African countries. Francois, Rainer and Trebbi (2012) show that, even in autocratic countries, representation is proportional to the size of the group, a finding which they interpret as the result of political instability. Given the risk of civil wars, political leaders

try to co-opt into government elites members of other ethnic groups. Bates (1983), Bratton and Van de Walle (1997) and Posner (2005) document the existence of patronage politics and the link between the ethnicity of the political leaders and public spending. Burgess, Jedwab, Miguel and Morjaria (2010) combine district-level panel data on road building in Kenya with historical data on the ethnicity and district of birth of political leaders to show that presidents disproportionately invest in their district of birth and those regions where their ethnicity is dominant. Similarly, Miguel and Zaidi (2003), Franck and Rainer (2012) and Moser (2008) show how African leaders use public expenditure to support individuals who come from their region of origin or share their ethnicity.

The paper is also linked to the literature on the persistence of institutions and the effect of institutions on development (Acemoglu, Johnson and Robinson, 2001, 2002; Rodrik, Subramanian and Trebbi, 2004, Alesina and Spoloare, 2003; Spoloare and Wacziarg, 2005). The analysis of limited national institutions, weak states and the inability to disseminate power is also central to Acemoglu (2005), Acemoglu, Ticchi and Vindigni (2011) and Besley and Persson (2010, 2011).

From a methodological point of view, the paper builds on the literature on matching models (Angrist and Pischke, 2008) and county-pair analysis (Dube, 2009; Naidu, 2010). With regard to development in Africa, this methodology has been pushed forward by Michalopoulos and Papaioannou (2012, 2014), who match partitioned groups in order to neutralize biases deriving from omitted variables related to ethnicity. In their analysis they focus on the effect of the rule of law and control of corruption within similar ethnic groups and find no significant effect of national institutions on development in Africa. By contrast to their analysis the current paper looks at the long-term effect on development of the Scramble for Africa itself beyond any sort of pre- and post-colonial institutional effects, which are ruled out by country and ethnic fixed effects.

3 Data Description

3.1 Data on Ethnic Groups

The Ethnographic Atlas (Murdock 1959; 1967) represents the main source of data; it provides information on economic activity and institutional organizations for 412 cultural clusters, representing 843 ethnic tribes in Africa[1]. For each group the Atlas reports the geographical coordinates and maps which have been digitized and made available by Nunn and Wantchekon (2011). This source is merged with spatial data on African administra-

[1]Among these 843 groups there are 8 uninhabited regions which will be dropped from the analysis

tive boundaries from GADM (Global Administrative Database) in order to map ethnic groups inside national boundaries. The intersection between ethnic location and national boundaries determines the partitioned groups.

Figure 1 shows all the possible partitioned groups within Africa. A group which has been partitioned is considered a new *ethno-country group* even though it shares the same culture, institutions and economic dependence as the original group. After considering all the possible partitions, the number of *ethno-country groups* in Africa increases to 1300, among which 830 groups have been affected by partition. Appendix 1 reports these groups and the number of countries between which the group has been partitioned. Of course it is possible that migration and the displacement of people after conflicts have changed the spatial distribution of groups, which may impose a limit to our analysis. However, statistical investigation by Michalopoulos and Papaioannou (2012) shows that such an effect is not particularly severe and the conjectured change of spatial distribution is not such as to affect the empirical analysis.

Figure 1: Ethnic and Partitioned Groups

Figure 2 shows the effect of the partition for a typical ethnic group. As a result of the Scramble for Africa, the Aulliminden group has been split into three new *ethno-country groups*, which we refer to as Mali-Aulliminden, Niger-Aulliminden and Algeria-Aulliminden. These groups occupy respectively 14.9 percent, 10.7 percent and 0.007 percent of the total country surface area. According to the existing literature (i.e. Alesina, Easterly and Matuszesky, 2010) these three groups should be affected by partition in the same way, given that it is the partition that matters. However, if we consider the hypothesis that the political, social, and cultural salience of the cleavage depends on the size of the group in relation to the size of the political and social arena in which the groups are located, then each of these three groups should be affected by the partition in a different way. In fact, the size of the Algeria-Aulliminden group is relatively insignificant with respect to the national political and social structure affecting the prospects of regional development for this group. As a consequence, the GDP per capita for the Algerian-Alluminden group should be much smaller than the GDP per capita for the other two Alluminden partitioned groups.

Figure 2: Partition of the Aulliminden Group

3.2 Data on GDP

Looking, as we do, at the level of development across groups implies that the measure of development must be at an ethnic group level. However, there are no sources that can provide such information directly. Therefore, in order to overcome this limitation we use estimates about total economic activity from Nighttime Lights satellite imagery provided by the NOAA/NGDC (Ghosh, Powell, Elvidge, Baugh, Sutton and Anderson, 2010). This source provides spatially disaggregated 1 km^2 data on total economic activity, which is recorded using a thirty two bit floating number (ranging from 0 to 147.682). In order to create spatially disaggregated data on economic activity the authors first estimate the total economic activity for each administrative unit by multiplying the sum of lights (i.e. the sum of brightness values of the lights for all lit areas) of each administrative unit by a coefficient obtained from regressing GDP (Gross Domestic Product) and GSP (Gross State Product) on lights[2]. Then they spatially distribute *"the estimated total economic activity of each administrative unit into 1 km^2 grid cells based on the percentage contribution of agriculture, the nighttime lights image, and the LandScan population grid"* (Ghosh, Powell, Elvidge, Baugh, Sutton and Anderson 2010, p. 151).

Using light density to obtain sub-national estimates of economic activity has been quite popular in the last few years (i.e. Henderson, Storeygard and Weil, 2012; Elvidge, Baugh, Kihn, Kroehl and Davis, 1997; Doll, Muller and Morley, 2006). Henderson, Storeygard and Weil (2012) use a similar approach to obtain estimates of GDP growth at sub-national level for 188 countries over 17 years. They use these estimates to evaluate whether, over the last 17 years, coastal areas have grown faster than non-coastal areas; whether primate cities have grown faster than hinterlands; and whether malarial areas have had a better growth experience than non-malarial areas[3]. They show that the implications of using sub-national estimates of GDP growth are quite different from standard results from the cross-country analysis (i.e Mellinger, Sachs and Gallup, 2000).

The spatial distribution of economic activity in Africa is shown in Figure 3. Darker areas denote regions of more intense economic activity while lighter areas denote regions with less or no economic activity (i.e. the Sahara Desert). The advantage of disaggregated data from an economic activity map is that it provides analytical flexibility given that

[2]See Chen and Nordhaus (2010) and Henderson, Storeygard and Weil (2012) for a more detailed discussions of regressions used to map lights into a proxy of GDP.

[3]One of the reasons why estimates of economic activity are normally preferred (rather than using straight light data) relates to measurement errors in light density related to cross-country cultural differences in the use of night-lights, gas flares, differences in lights sensitivity across satellites, blooming and bleeding, attenuation of lights for areas with low economic activity, etc. (Chen and Nordhaus, 2010; Henderson, Storeygard and Weil, 2012)

data can be aggregated to units of different sizes. As a consequence we can use these disaggregated data in order to construct proxies of development for each ethnic group in our sample[4].

Figure 3: $1km^2$ Data on Total Economic Activity (converted into a Shapefile)

In order to verify the reliability of our proxy for economic activity[5] we compare our data with a measure of mean light density from Michalopoulos and Papaioannou (2013) for the 679 observations in their sample. Table 1 shows the pairwise correlation which is close to 0.89 providing enough evidence on the validity of our proxy for GDP per capita.

Table 1: Pairwise Correlation between Economic Activity and Light Density

4 Partition and Development

4.1 Empirical Strategy

We start our investigation using a cross-group analysis in order to estimate the effect of the partition on development. The basic model estimated to evaluate such an effect can be written as:

$$y_{i,c} = \delta_c + \beta_1 Partition_{i,c} + \beta_2 Grp\_Share_{i,c} + \beta_3 X_{i,c} + \epsilon_{i,c} \quad (1)$$

where $y_{i,c}$ is our proxy for mean economic activity (GDP) for group i in country C; δ_c denotes country-specific effects; $Partition_{i,c}$ is the dummy for whether the group has been partitioned or not; $Grp\_Share_{i,c}$ represents the population share of the group i in country C; and $X_{i,c}$ is a set of control variables. The error $\epsilon_{i,c}$ is double clustered in order to capture potential auto-correlation within ethnicities and countries.

Then we restrict our analysis to partitioned groups only, controlling for ethnic fixed effects in order to flush out biases related to ethnic characteristics. Therefore the model to be estimated in this case can be written as:

[4]The proxy for GDP is constructed using the Zonal Statistics in Qgis 2.01

[5]The mean economic activity represents the mean value at a $1km^2$ cell within a polygon (ethnic tribe). Intuitively this is nothing else that the total economic activity within a polygon (ethnic tribe) divided by the total number of $1km^2$ cells within the polygon.

$$y_{i,e,c} = \delta_e + \beta_1 Grp\_Share_{i,e,c} + \beta_2 X_{i,e,c} + \epsilon_{i,e,c} \quad (2)$$

where δ_e now captures ethnic-specific effects and $y_{i,e,c}$ is a measure of development for the partitioned group i belonging to the ethnic group e within country C. The inclusion of ethnic fixed effects allows us to deal with omitted variables related to group-specific characteristics (i.e. pre-colonial institutions, pre-colonial development, etc.) which in some way may be correlated with the share of partitioned groups. From a certain point of view the estimator compares measures of development for group i in country C_1 with exactly the same group i but in country C_2. Therefore groups are matched on the basis of similar unobservable characteristics which may affect the estimates.

The variable of interest is $Grp\_Share_{i,c}$. If the partition matters because of the creation of relatively small groups then we should expect a significant effect on development of the population share of the group ($Grp\_Share_{i,c}$). However, if it is the partition itself that matters, then the population share of the resulting group should not matter, given that the groups resulting from the partition should be affected in a similar way independently of their size (same as in Alesina, Easterly and Matuszesky, 2010). Of course, the case in which both effects (partition and group share) matter is also possible.

Table 2 provides descriptive statistics for some of the main variables in our model. Almost 64 percent of the groups in our dataset have been affected by the partition (Partition Dummy). The average population share of a typical partitioned group is smaller than 4 percent of the country's population. The surface area of the typical partitioned group (as a share of the country's surface area) is also approximately equal to 4 percent. In fact the correlation between the share of the surface area and the population share of the group is almost 0.75. A typical partitioned group has an average population of almost 613,000. At the same time the average population in the region that is included within the ancestral ethnic homeland is almost equal to 1 million. As a consequence the population share of the average partitioned group as a share of the population within the ethnic homeland is around 64 percent. Even in this case the population share is closely correlated with the share of the surface area of the partitioned group (as a share of the ethnic homeland surface area) with a correlation of 0.95. The mean economic activity (GDP) is 0.79 and the population density is almost 51 inhabitants per 2.5 by 2.5 arc-minutes (approximately 25 sq km at the Equator)[6]. Following Michalopoulos and Papaioannou (2012) we transform the dependent variable by adding a small number (0.01) to economic activity before taking

[6]Observations with a population density, GDP and group share equal to 0 are dropped from the analysis. For this reason in the following tables the number of observations goes down.

logs in order to correct for the fact that the distribution is skewed toward zero. However we will also report estimates for the log of the mean economic activity without correcting for such a small number.

Table 2: Descriptive Statistics

4.2 Results

Table 3 shows the results from the cross-section analysis in which all groups (partitioned and not) are pooled together in order to evaluate whether partition matters. In the first model (Model 1), we control only for country fixed effects and we find a positive and significant effect of the population share of partitioned groups ($Grp\_Share_{i,c}$) on development, which increases income by almost 1.6 percent per one standard deviation in the share of the group. The dummy for whether the group has been partitioned has also a significant effect on development, reducing GDP per capita by almost 11 percent. In Model 2 we include controls for the population share of the group, population density estimated (from the Gridded Population of the World) and a set of geographical controls. Once controlling for a full set of controls, the effect of the partition turns to be insignificant while the proxy for the share of the ethnic groups still exerts a positive effect on development. In Model 3 we enter additional controls for types of settlement, dependence on gathering, and jurisdictional hierarchy beyond the local community, which Michalopoulos and Papaioannou (2013) find to be closely associated with regional development. Murdock (1959) represents the source for these additional controls. Even after entering these additional controls, the results are almost unchanged. There is still an insignificant effect of the partition on development, while the population share of the group ($Grp\_Share_{i,c}$) exerts a significant and positive effect on development.

Table 3: Cross-Group Analysis

Given that the cross-group analysis is likely to be affected by omitted variables correlated to ethnicity in the following specifications, we restrict our analysis to within partitioned ethnic homelands in order to have more precise estimates of the effect of the share of the group. Specifically, we confine estimates to the same groups residing on both sides of state borders, controlling for ethnic-fixed effects. Matching groups which share the same

unobservable characteristics and the inclusion of ethnic fixed effects allows us to compare the level of development for the same group on both sides of the border, which from a practical point of view means that the estimator compares measures of development for group i in country C_1 with exactly the same group i in country C_2. Therefore the estimator will represent a sort of quasi-experiment and because of that it will flush away all the ethnic characteristics that might affect estimates.

The identification of the effect is conditional on the fact that state boundaries are randomly drawn. Therefore a possible threat to the identification may be related to the fact that small partitioned groups may have been historically marginalized either in terms of historical development or geographical factors. In this case the two (or more) partitions of each ethnicity may not be appropriate counterfactuals. To test whether there are any systematic differences between small and large partitioned groups in Table 4 we estimate simple ethnicity fixed-effects regressions associating the variables that reflect geography and pre-colonial development in the share of partitioned groups. Table 4 reports the coefficients on the share of the partitioned groups for each specification. Differences in pre-colonial development (proxied by population density in 1800) do not systematically differ between groups (Model 1). Differences in terms of geography and disease environment (proxied by malaria suitability) are also not significantly different (Models 2-5). The distribution of natural resources is also not systematically linked to the size of partitioned groups (Model 6). Therefore the lack of systematic differences between small and large groups seems to support the identification design.

Table 4: Group Share and Pre-colonial/Geographical Characteristics

When we confine estimates to partitioned groups only (Table 5) the coefficient on the population share of the partitioned group ($Grp\_Share_{i,c}$) is still significant at a 1 percent level and the coefficient remains almost unchanged, with the population share of the partitioned group increasing income by almost 1.4 percent per one standard deviation in the former (Model 1). In Model 2 we include country fixed effects in order to rule out country characteristics (i.e. institutions, national fractionalization, etc.) and the coefficient on the share of the partitioned group increases to 1.6. Spillover effects related to the spatial distribution of groups sharing the same ethnicity may affect the error term, which is therefore likely to be spatially correlated. Because of this problem with the efficiency of the estimator in Model 3 and Model 4 we use Conleys spatial HAC estimator (2008) to adjust OLS standard errors for spatial correlation. For each model we use a

different distance threshold in order to test the robustness of the results. In Model 3 we use a distance threshold equal to 100km and in Model 4 we increase the threshold to 200km. Standard errors after controlling for spatial correlation decrease by almost one-half. As a result, the effect of the size of partitioned groups is still significant at a 1 percent level. In addition, increasing the distance threshold from 100km (Model 3) to 200km (Model 4) does not seem to have any great effect on standard errors. In Table A1 in the Table Appendix we re-estimate Model 2 using a quadratic form on the population share of the partitioned group in order to control for diminishing returns given that the effect may increase till the share becomes close to 50%(majority of seats in parliament/government) and then turns either negative or insignificant. The squared form is significant however the effect seems to be maximized when the share of the group is close to 58%. In our sample there are only few groups with a share larger than 58% and because of that the inverted U-shape relation may be the result of these outliers.

Table 5: Matched-Groups

Because the small size of groups may prevent firms and farmholders from taking advantage of scale economies in production affecting production costs, productivity and potential for growth, in Table 6 we enter additional controls for the log of the group total surface area (in m^2) (Model 1). In Model 2 we also control for the total population within the group. Though the two variables are closely correlated with the population share of the partitioned group the latter still has a significant effect on GDP. A one standard deviation in the share of the partitioned group increases income by almost 0.11 percent. Finally in Model 3 we change the functional form of the dependent variable[7] and the results still hold. The estimated effect increases to almost 2 percent for a one percentage change. This larger effect is likely to reflect the fact that the distribution is skewed toward zero, increasing the total variance, which justifies our approach of adding a small quantity to the log in order to normalize the distribution.

Table 6: Controlling for Surface Area, Total Population, and Alternative Functional form on GDP

Michalopoulos and Papaioannou (2014) argue that problems in the definition of ethnic boundaries in Murdock (1959, 1967) can affect estimates and hence groups which are

[7]Instead of the log of (0.01 + GDP) we use the log of GDP

relatively small should not be used when evaluating levels of development across partitioned groups. In order to deal with this problem they restrict their analysis to groups with at least 10 percent of their ethnic homeland belonging to more than one country. Consequently in Table 7 we confine estimates first to groups with at least 5 percent and then to groups with at least 10 percent of the ethnic homeland belonging to more than one country.

In Model 1 we consider first groups with at least 5 percent of the ethnic homeland belonging to more than one country. The number of observations in Model 1 drops to 571, but the effect of the share of groups is still significant at a 1 percent level with an effect on development almost equal to 1.5 percent per a one percent change in the share of the group. Then in Model 2 we go on to consider groups with at least 10 percent of their ethnic homeland belonging to more than one country. In this case the number of observations drops further, to 483, whereas the coefficient on the size of the group increases quite significantly (to 2.2 percent) with the effect still significant at 1 percent.

Table 7: Dropping Small Groups

5 Robustness Checks

In order to provide evidence of the robustness of our results we carry out a series of robustness checks in order to make sure that the effect of the population share of partitioned groups still retains its significance. First we try to re-estimate the effect of the share of the partitioned group on development using an IV estimator. Then we check whether the results still hold using an RD approach. Finally we replace our dependent variable for luminosity with another dependent variable which is likely to be associated with regional development.

5.1 Regression Discontinuity

We start our robustness checks using a regressions discontinuity approach (Table 8) which identifies the effect of the share of groups at the border. In our case the running variable is the distance from the centroid of an ethnic area to the national border. Following Imbens and Lemieux (2008), Lee and Lemieux (2010) we present results for specifications where we use a cubic RD-polynomial in distance to the border (Model 1) and for a local

linear regression (Model 2 and Model 3) where we confine estimates to groups within *100 km* and *50 km* from the border respectively. Table 9 reports our estimates. In Model 1 our proxy for the population share of partitioned groups is still significant at a 5% level exerting an effect on development which is close to 1.5 percent for a percentage change in the share. In Model 2 we confine estimates to groups within *100 km* from the border in order to estimate the local average effect. The number of observations drops to 728 but the effect of the share of partitioned groups is still significant at a 5% level. Finally in Model 3 we confine estimates to groups within *50 km* from the border. The sample drops further to 563 observations and the estimated effect is not significant. However, this insignificant effect is to be imputed to the loss of efficiency of the estimator due to the smaller number of observations. In fact the point estimate is almost the same as the one estimated in Model 1, although the standard error increases quite significantly (from 0.48 to 0.99) which reflects the fact that degrees of freedom decrease quite significantly. There is still a significant economic effect of the population share of partitioned group on development, though the statistical effect is not significant because of the loss of efficiency of the estimator.

Table 8: Regression Discontinuity

5.2 Instrumental Variable

As an additional test we use an IV estimator which takes the relative size of the group (i.e. surface area of the group/country surface area) as an instrument for estimating the population share of the partitioned group. The relative size of the ethnic group seems to be a natural instrument for the population share, given that the latter is likely to be larger for larger groups. In fact the correlation between the two variables is more than 0.75. The close correlation between the size and population share of the partitioned group is confirmed by the first-stage statistics reported in Table 9. The size of the partitioned groups has a strong and significant effect on the population share of the partitioned group, with the latter increasing by almost 0.8 percent per every one percent change in the relative size of the partitioned group. As a result, the Cragg-Donald F-statistics is much larger than the Stock and Yogo critical values, avoiding the problems related to biases in the distribution. With regard to the second-stage statistics, in Model 1 we confine estimates to all partitioned groups and the effect of the population share is still significant at a 1 percent level with a coefficient which is close to the one estimated using regression discontinuity.

In Model 2 and Model 3 we restrict the sample to the partitioned groups representing either at least 5 percent or at least 10 percent of the ethnic homeland and the effect of the population share of the partitioned groups is still significant at a 10 percent level at least. In addition, consistent with the OLS results, the coefficient increases quite significantly when we confine estimates to partitioned groups representing at least 10 percent of the ethnic homeland.

Table 9: IV Estimates

One of the potential concerns with the IV results above is related to the fact that the size of the group may not be completely exogenous, given the effect of the relative size of the group on economies of scale, natural resources, access to the sea, etc. To explore the sensitivity of our results to different degrees of violation of the exclusion restriction, we follow the methods proposed by Conley, Hansen and Rossi (2012). The idea behind this sensitivity analysis is to relax the assumption that the direct effect of the instrument on the dependent variable is equal to zero using different priors for the direct effect of Z on Y. Therefore, in order to test the sensitivity of the IV results, we need to choose a set of different priors (δ) about the potential direct effect of the instrument. To choose the set of different δs we regress the instrument and the endogenous variable on the dependent variable (controlling for same variables as above) in order to estimate the potential direct impact of the instrument on the dependent variable. Then we choose a quite large interval around this estimated effect in order to account for the potential bias of OLS estimates. The direct effect of the instrument on GDP per capita is negative but not significant (Table A2 in the Appendix). Given the insignificant effect we consider $\delta \in (-0.8, 0.8)$ as a potential interval of priors where the lower-bound is ten times larger than the insignificant estimated effect in Table A2. Results of the sensitivity analysis are reported in Figure 4. Dashed lines present the union of 2SLS 95% confidence intervals around the IV coefficient, using Conley, Hansen, and Rossi (2012) UCI approach. Our IV estimates hold statistically significant even with substantial departures from the assumption that the direct effect of the instrument is identically zero. Confidence intervals include zero only for quite large deviations from the exclusion restriction (δs larger than 0.6) which we consider highly unlikely. In addition, confidence intervals are relatively stable denoting the strength of the instrument.

The result of the sensitivity test above together with the fact that the estimated effect from using an IV is quite similar to the estimated effect from the regressions discontinuity

approach above seems to provide some support to the idea that the ratio of the surface area of the group to the country surface area may satisfy the exogeneity condition.

Figure 4: Sensitivity to Violations of the Exclusion Restriction

5.3 Alternative Dependent Variables

The last test consists in replacing the dependent variable with household data from the Afrobarometer survey to evaluate whether results still hold when we use a different outcome variable which captures income at an individual level. Appendix A discusses the approach we use in order to merge ethnic groups from Murdock (1959) with language data from the Afrobarometer IV[8]. In Model 1 we regress our variable for the population share of the group on a variable proxying how often the individual has gone without cash income (Q8E) controlling for country and ethnic fixed effects and a full set of geographical, district, and individual controls. The variable ranges from *never* (equal to 1) to *always* (equal to 5). In Model 2, we replace Q8E with Q8D (how often the individual has gone without cooking fuel). We choose these two variables because we believe that they are closely correlated to income given that they are directly related to cash income. For both models we find a significant and negative effect of the share of the group suggesting that the probability of going without cash and fuel is much lower for individuals within larger groups (Table 10). As a result, the micro-based results with the Afrobarometer data suggest that our benchmark estimates with satellite light density as the dependent variable are not an artefact of the luminosity data.

Table 10: Alternative Dependent Variables

6 Channels

We now turn to the analysis of potential channels through which the share of partitioned groups may affect development. In principle there are several competing theories of why the share of ethnic groups may matter for development. We group these theories in three categories.

.

[8]Unlike the third wave, which provides information on ethnicity, the Afrobarometer IV provides information only on the languages of the respondents

6.1 Ethnic Voting, Political Power and Ethnic Patronage

The share of the group may affect political representation, the provision of local public goods, investment in infrastructure and therefore regional development. In fact, a number of studies show a close correlation between ethnic identity voting, representation, and ethnic patronage. Eifert, Miguel, and Posner (2010) show that the probability of individuals identifying themselves in ethnic terms increases for every month closer the country is to a competitive presidential election. Casey (2012) shows that in the 2007 parliamentary elections the Sierra Leone Peoples Party (SLPP), which has close ties with the Mende group, captured 24 of the 25 seats in the South (homeland of the Mende). At the same time, the All Peoples Congress (APC), which is linked to the Temne group, won 36 out of 39 seats in the Northern Province (homeland of the Temne). Posner (2005) provides evidence of the close ties between ethnic identity and voting in Zambia, while Long (2012) shows a similar relationship between ethnic identity and party choice in Kenya. The link between ethnic identity and voting affects representation in parliament and the likelihood of cabinet office (Rainer and Trebbi, 2014; Francois, Rainer and Trebbi, 2012) and therefore ethnic patronage. In fact, there is evidence that ethnically based parties tend to redistribute toward their ethnic group and citizens tend to vote for candidates who represent their group regardless of their quality. Evidence on ethnic political patronage is provided by Burgess, Jedwab, Miguel and Morjaria (2015) who find that a large share of road investments in Kenya may be explained by political appointments and ethno-favouritism in Kenyan politics. Kramon and Posner (2012) look at education in Kenya. Using data on the educational attainment of more than fifty thousand Kenyans dating back to the colonial era and information about the ethnic identities of Kenyan presidents and cabinet members, they find that having a co-ethnic as president during one's primary school-age years is associated with about a one-quarter of a years increase in years of primary schooling and substantial increases in the probability of attending and completing both primary and secondary school. Franck and Rainer (2012) use data from 18 countries to show that national leaders have significantly increased primary school attendance and reduced infant mortality within their ethnic group. Hodler and Raschky (2011) use light images from satellites to see whether national leaders favour their birth region when they take power and whether this effect is more muted under democracy or autocracy. The problem with ethnic patronage is that the quality of politicians is likely to be affected when the impossibility of holding politicians to account undermines the quality of political candidates, which then leads to undesirable governance outcomes such as corruption (Banerjee and Pande, 2007). As a result the prospects of development are undermined.

In order to provide basic evidence on whether the share of the group affects the local provision of public goods and regional investment in infrastructures we use spatial data on roads and railroads from the FAO-GEONETWORK and we construct a measure for the distance (in metres) of roads and railroads from the partitioned group. The idea is that if ethnic patronage matters then small groups which lack political representation should receive less investment and, therefore, they should be further from roads and railwaysOne of the problem with our variable for roads is that the data include primary, secondary and tertiary roads which in some ways is likely to bias downward the effect of the share of group on the distance from roads given the relative insignificant effect (and the relative wide coverage) of tertiary roads on transportation costs and development..

To estimate the effect of the share of the group on roads and railroads we use a spatial HAC estimator in order to take into account the large spatial autocorrelation in roads and railroads networks. Table 11 presents results which use exactly the same specification (and controls) as previous models do. Of course, the results are unlikely to be causal given that we do not exploit any natural experiment. However, the estimated effects point in the right direction. In Model 1 we show the effect on the distance from roads (in log) and we find that a one percent increase in the share of the group decreases the distance from roads by almost 4.7%. In Model 2 we show estimates for the distance from railroads. The effect of the share of the partitioned group is still significant at a 1% level and the effect of the share of the group on the distance is negative. A one percent increase in the share of the partitioned group reduces the distance from roads by almost 7.3%.

Table 11: Partitioned Groups and Infrastructures

6.2 Dual Legal System, Persistence of Informal Institutions, and Trust

Smaller groups may, however, rely more on customary institutions affecting property rights, the rule of law, land allocation and therefore development. For example, Mamdani (1996) argues that the legacy of colonialism in Africa is a *"bifurcated state"* where decentralized despotism rules in the rural areas and democratic institutions govern relations in the urban areas. The creation of customary institutions in the rural areas empowers local chiefs/tribal leaders and their influence persists nowadays, affecting property rights (Bubb, 2013), the redistribution of land (Goldstein and Udry, 2008), corruption and communal despotism (Lange, 2004).

In order to test the idea that smaller groups are more likely to rely on informal institutions we use individual data from the Afrobarometer IV. In Model 1 we regress the population share of the partitioned group on a dummy variable coded one if the tribal leader has the primary responsibility for collecting income taxes (Q58D). In Model 2, we replace the dependent variable with a dummy for whether the tribal leader has the primary responsibility for solving local disputes (Q58E). Finally, in the last two models we consider whether tribal leaders have the primary responsibility for maintaining law and order (Q58H) and for allocating land (Q58F). We find a significant effect of the share of the group on proxies for whether the tribal leader has the primary responsibility for maintaining law and order (Model 3) and for the allocation of land (Model 4). In fact, larger groups tend to rely less on tribal leaders with regard to these two dimensions. The effect on the collection of income taxes is also negative, but not significant (Model 1). The effect on solving local disputes is also not significant (Model 2).

Table 12: Partitioned Groups and Informal Institutions

When formal institutions fail, networks of socially connected individuals can increase economic efficiency (Mundi, 2014), leading to a sort of second best[9]. Network relationships largely depend on trust given that individuals who trust less will be connected with fewer nodes. As a result groups which largely rely on informal institutions may have a higher level of trust since informal institutions tend to rely on human interactions. Therefore, given the effect of the share of the group on informal institutions we should also observe a significant relationship between the share of groups and trust. Specifically we should expect a negative effect of the share of ethnic groups on trust.

To evaluate the effect of the share of groups on trust we consider three questions from the Afrobarometer IV. The questions ask whether the respondent trusts relatives (Q84A), trusts other people he knows (Q84B), and trusts individuals from other groups (Q84C)[10]. The variables range from *not at all* (equal to 1) to *a lot* (equal to 4). In Model 1 we do not find any significant effect of the share of partitioned groups on the level of trust in relatives. However the level of trust in other people (Model 2) and the level of trust in individuals from other groups (Model 3) is significantly affected by the population share

[9]Fafchamp (2001) shows that network externalities may steer potential investors towards sectors of activity where they can benefit from these externalities, while Fafchamp and Lund (2003) show a positive effect of network relationships on risk sharing in the rural Philippines.

[10]Michalopoulos and Papaioannou (2014b) use same variables to look at the various historical and contemporary functions of tribal leaders (chiefs) and illustrate their influence on various aspects of the economy and the polity

of the group. In fact, larger groups tend to have a lower level of trust in other individuals (either from the same or different groups). These results are consistent with the idea that groups which tend to rely more on informa institutions tend to trust more given that they tend to rely more on repeated social interactions. Of course, this is not a first best, but given the evidence on trust and economic growth one can consider the effect of the share of groups on trust as a sort of second best.

Table 13: Partitioned Groups and Trust

Therefore, to sum up, we find that the level of development of small partitioned groups may be adversely affected because of ethnic patronage and the persistence of informal institutions. Still, the effect on communality networks and trust is not likely to hamper the level of economic development of these groups and should actually have a positive effect on the level of development. Of course, these results cannot be interpreted as causal in nature, but they provide some basic evidence related to factors which may explain the relationship between the share of a group and development.

7 Conclusions

The economic and political science literature has always maintained that the arbitrary imposition of state boundaries is one of several factors that explain the poor development in Africa. However the analysis in this paper seems to suggest that it is not the partition itself that matters for development but the way in which groups have been partitioned. To be precise, there is a significant effect only in cases where the partition creates small ethnic groups which lack political representation. From this point of view, a more inclusive political system could be beneficial in reducing such an effect.

APPENDIX
Merge with the Afrobarometer

The Afrobarometer (2008) Fourth Round is our main source for formal and informal institutions[11]. Different from the Third Round (Afrobarometer, 2005) which provides data on the ethnicity of individuals (variable Q79), the Fourth Round does not provide a variable which directly indicates the ethnicity of individuals. However, each individual in the survey is asked to report his native language. Therefore in order to match individuals in the Afrobarometer with data on ethnic groups in Murdock (1959, 1967) we rely on information on native languages[12]. For each individual in a country-region-district we check which ethnic group in such country-region-district speaks such a language and then we match these individuals with ethnic groups in Murdock. Of course, there are practical issues related to the fact that languages in the Afrobarometer do not always match with names of ethnic groups in Murdock. Therefore to understand which ethnic group in a given country-region-district speaks a given language we rely on information from the Ethnologue and from the Joshua Project.

There are three main indicators of *"language/ethnicity"* in the Fourth Round which matter for our analysis. The first one is the language of the respondent (variable Q3); the second indicator is the language of the interview (Q103); and the third indicator is a question related to the spoken languages (Q88E). We use the information from these variables together with the data on country (COUNTRY) and regional bases of each group (REGION and DISTRICT) in order to merge the data on ethnicity from the Afrobarometer with the data on ethnic groups from Murdock. We first try to match the language of the respondent (variable Q3) with Murdock's data on ethnicity, though this is not always straightforward. In fact, in some cases the reported language is French, English or Portuguese. For those individuals who report a European language as a spoken language we then check if the interviewer reports the language in which the interview is conducted (variable Q103) and if this language is different from French, English or Portuguese. If the language in which the interview is conducted is different from the three European languages above then we use this additional information to match data. If the language of the interview is not reported (or not different from English, French or Portuguese),

[11]The 20 countries covered by the fourth round (Afrobarometer, 2008) are the following: Benin, Botswana, Burkina Faso, Cape Verde , Ghana, Kenya, Liberia, Lesotho, Madagascar, Malawi, Mali, Mozambique, Namibia, Nigeria, Senegal, South Africa, Tanzania, Uganda, Zambia, Zimbabwe

[12]Nunn and Wantchekon (2011) use the same matching strategy. The only difference is that they use data on ethnicity of individuals (Q79) which is available for the Third Round but not available for the Fourth Round.

then we finally look at the spoken languages and we merge the spoken language with the related ethnicity (this is done for fewer than 50 obs.) We assume that individuals within a country-region-district speaking the same language belong to the same ethnic group.

References

[1] Acemoglu, Daron, (2005): Politics and Economics in Weak and Strong States, Journal of Monetary Economics, vol. 52(7), pages 1199-1226

[2] Acemoglu, D. and S. Johnson (2005): Unbundling Institutions, Journal of Political Economy, 113(5), 949995.

[3] Acemoglu, Daron, (2005): Politics and Economics in Weak and Strong States, Journal of Monetary Economics, vol. 52(7), pages 1199-1226

[4] Acemoglu, D., S. Johnson and J. A. Robinson (2001): The Colonial Origins of Comparative Development: An Empirical Investigation, American Economic Review, 91(5), 13691401.

[5] Acemoglu, D., S. Johnson and J. A. Robinson (2002): Reversal Of Fortune: Geography And Institutions In The Making Of The Modern World Income Distribution, Quarterly Journal of Economics, 107(4), 12311294.

[6] Acemoglu, D., S. Johnson and J. A. Robinson (2005): Institutions as a Fundamental Cause of Long-Run Growth, in Handbook of Economic Growth, ed. P. Aghion and S. N. Durlauf, pp. 109139. Elsevier North Holland, Amsterdam, The Netherlands.

[7] Acemoglu, D., J. A. Robinson (2012): Why Nations Fail. The Origins of Power, Prosperity and Poverty. Crown Publishers, New York, NY.

[8] Acemoglu, D., D. Ticchi and A. Vindigni (2011): Emergence And Persistence Of Inefficient States Journal of the European Economic Association, 9(2): 177-208

[9] Adewumi, J.B. J Egwurube. (1980): Role of traditional rulers in historical perspective. In Aborisade, Oladimeji (Ed.). Local Government and the traditional rulers in Nigeria. Ile-Ife: University of Ife Press.

[10] Ajala, Adekunle (1983): The Nature of African Boundaries; Afrika Spectrum, 18, 177-188

[11] Alesina, Alberto, Arnaud Devleeschauwer, William Easterly, Sergio Kurlat and Romain Wacziarg (2003): Fractionalization, NBER Working Papers 9411

[12] Alesina, A. and E. Spolaore (2003): The Size of Nations. MIT Press.

[13] Alesina, A., W. Eastely and J. Matuszesky, (2011): Artificial States, Journal of the European Economic Association, 9(2), 246277.

[14] Angrist, J. and J.-S. Pischke (2008): Mostly Harmless Econometrics. Princeton University Press, Princeton, NJ

[15] Asiwaju, A. I. (1985): The Conceptual Framework, in A. I. Asiwaju (Ed.), Partitioned Africans (pp. 1-18). New York: St. Martins.

[16] Bach, Daniel (1999): Regionalisation in Africa: Integration and Disintegration. Bloomington: Indiana University Press

[17] Banerjee A. and R. Pande (2007) Parochial politics: Ethnic preferences and political corruption, CEPR discussion paper no. 6381.

[18] Barbour, K. M. (1961): A Geographical Analysis of Boundaries in Inter-Tropical Africa. In K. M. Barbour R. M. Prothero (Eds.), Essays on African population (pp. 303-323). London: Routledge Kegan Paul

[19] Bates, Robert. (1983): "Modernization, Ethnic Competition, and the Rationality of Politics in Contemporary Africa." In States versus Ethnic Claims: African Policy Dilemmas, ed. Donald Rothchild and Victor Olunsorola. Boulder, CO: Westview Press

[20] Bayart, Jean-Franois. (1996): LHistoricit de lEtat Import, In J. Bayart (Ed.), La Greffe de lEtat (pp. 11-39). Paris: Khartala

[21] Bello, Adebayo (1995): The Boundaries Must Change, West Africa, p. 546

[22] Besley, T. and T. Persson (2010): State Capacity, Conflict and Development, Econometrica, 78(1), 134.

[23] Besley, T. and T. Persson (2011): Pillars of Prosperity. The Political Economics of Development Clusters. Princeton University Press, Princeton, NJ.

[24] Bratton, Michael, and Nicolas Van de Walle (1997): "Democratic Experiments in Africa: Regime Transitions in Comparative Perspective." New York: Cambridge University Press

[25] Brinks, Daniel (2003): Informal Institutions and the Rule of Law: The Judicial Response to State Killings in Buenos Aires and So Paulo in the 1990s. Comparative Politics 36 (1): 119.

[26] Brownlie, Ian (1979): African Boundaries: A Legal and Diplomatic Encyclopedia, London: C. Hurst

[27] Bubb, R. (2013): The Evolution of Property Rights: State Law or Informal Norms?, Journal of Law and Economics, 56(3), 555594.

[28] Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padr i Miquel (2015): "The Value of Democracy: Evidence from Road Building in Kenya." American Economic Review, 105(6): 1817-51

[29] Casey, K. (2012): Crossing Party Lines: The Effects of Information on Redistributive Politics, Stanford GSB Working Paper No. 2099

[30] Cederman, L.E., Wimmer, A. and Min, B. (2010): Why Do Ethnic Groups Rebel? New Data and Analysis, World Politics 62(1): 87-119

[31] Chen, Xi, and Nordhaus, Williams D. (2010): The Value of Luminosity Data as a Proxy for Economic Statistics, NBER Working Paper 16317

[32] Clapham, Christopher (1996b): Boundary and Territory in the Horn of Africa. In Paul Nugent A. I. Asiwaju (Eds.), African boundaries (pp. 237-250). London: Pinter

[33] Conley, Timothy G. (2008): "Spatial Econometrics." The New Palgrave Dictionary of Economics. Second Edition. Eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan.

[34] Conley, T. G., Hansen, C. B., Rossi, P. E. (2012). Plausibly exogenous. Review of Economics and Statistics, 94(1), 260-272

[35] Davidson, Basil (1992): The Black Mans Burden: Africa and the Curse of the Nation-State, New York: Times Book

[36] Doll, C. N., J.-P. Muller and J. G. Morley (2006): Mapping Regional Economic Activity from Night-Time Light Satellite Imagery, Ecological Economics, 57, 7592.

[37] Dowden, R. (2008): Africa: Altered States, Ordinary Miracle Portobello Books Ltd

[38] Dube, A., Lester, T. W. and Reich, M. (2009) Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties, Institute for Research on Labor and Employment Working Paper, 157:07.

[39] Easterly, William and Levine, Ross (1997): Africa's Growth Tragedy: Policies and Ethnic Divisions, The Quarterly Journal of Economics, vol. 112(4): 1203-50

[40] Easterly, William Levine, Ross (2003): Tropics, Germs and Crops: How Endowments Influence Economic Development, Journal of Monetary Economics, Elsevier, vol. 50(1), pages 3-39

[41] Eifert, B., E. Miguel, and D. Posner (2010): Political Competition and Ethnic Identification in Africa, American Journal of Political Science, 54(1), 495510

[42] Elvidge, C., K. Baugh, E. Kihn, H. Kroehl and E. Davis (1997): Mapping of City Lights Using DMSP Operational Linescan System data, Photogrammetric Engineering and Remote Sensing, 63, 727734.

[43] Englebert, P. (2000a): Solving the Mystery of the Africa Dummy World Development, 28 (10):1821-1835

[44] Englebert, Pierre (2000b): Pre-Colonial Institutions, Post-Colonial States and Economic Development in Tropical Africa, Political Research Quarterly 53: 7-36

[45] Englebert, P., S. Tarango and M. Carter (2002): Dismemberment and Suffocation: A Contribution to the Debate on African Boundaries, Comparative Political Studies, 35(10), 10931118.

[46] Esteban, Joan and Debraj Ray (2008): Polarization, Fractionalization and Conflict, Journal of Peace Research, 45,163-182.

[47] Ethnologue (2005): Languages of the World, SIL International; Fifteenth edition.

[48] Fafchamps, Marcel. (2001) "Networks, Communities and Markets in Sub-Saharan Africa Implications for Firm Growth and Investment," Journal of African Economies 10: 109-142.

[49] Fafchamps, Marcel, and Susan Lund (2003). Risk Sharing Networks in Rural Philippines. Journal of Development Economics 71(2): 26187

[50] Fearon, James D (2005): Ethnic Mobilization and Ethnic Violence In Barry R. Weingast and Donald Wittman, Eds., Oxford Handbook of Political Economy. Oxford: Oxford University Press

[51] Fearon, James D (2003): Ethnic and Cultural Diversity by Country, Journal of Economic Growth, Springer, vol. 8(2), pages 195-222

[52] Fearon, James and David D. Laitin (2003): Ethnicity, Insurgency and Civil War. American Political Science Review 97(1): pp.75-90

[53] Francois, Patrick, Ilia Rainer and Francesco Trebbi (2012): How Is Power Shared In Africa?, NBER Working Paper No. 18425

[54] Franck, Raphael, and Ilia Rainer. (2012): "Does the Leader's Ethnicity Matter? Ethnic-Favoritism, Education and Health in Sub-Saharan Africa." American Political Science Review, 106(2), pp 294-325

[55] Gennaioli, N. and I. Rainer (2006): Precolonial Centralization and Institutional Quality in Africa, in Institutions and Norms in Economic Development, by M. Gradstein and K. Conrad, (Eds.) MIT Press

[56] Glaeser, E.L., La Porta R., de-Silanes F. L. and Shleifer A. (2004): Do Institutions Cause Growth? Journal of Economic Growth, Springer, vol. 9(3), pp. 271-303,

[57] Ghosh, T., Powell, R., Elvidge, C. D., Baugh, K. E., Sutton, P. C. Anderson, S. (2010): Shedding Light on the Global Distribution of economic Activity The Open Geography Journal (3), 148-161

[58] Goldstein, M., and C. Udry (2008): The Profits of Power: Land Rights and Agricultural Investment in Ghana, Journal of Political Economy, 116(6), 9811022.

[59] Griffiths, Ieuan. (1996): Permeable Boundaries in Africa, In Paul Nugent A. I. Asiwaju (Eds.), African Boundaries (pp. 68-83). London: Pinter

[60] Helmke G. and Levitsky S. (2004): Informal Institutions and Comparative Politics: A Research Agenda Perspectives on Politics, 2(4): 725-40

[61] Henderson, V. J., A. Storeygard and D. N.Weil (2012): Measuring Economic Growth from Outer Space, American Economic Review, 102(2): 994-1028

[62] Herbst, J. (2000): States and Power in Africa. Princeton University Press, Princeton, NJ.

[63] Hodler, R. and Raschky, P. A. (2010). Foreign Aid and Enlightened Leaders. Working Papers 10.05, Swiss National Bank, Study Center Gerzensee.

[64] Horowitz, Donald L. (1985): Ethnic Groups in Conflict, University of California Press

[65] Hsiang, S., 2010, Temperatures and cyclones strongly associated with economic production in the Caribbean and Central America. Proceedings of the National Academy of Sciences, 107(35): 1536715372

[66] Huntington, S. P. (1991): The Third Wave: Democratization in the Late Twentieth Century, University of Oklahoma Press, 1991

[67] Imbens, G. W., and T. Lemieux (2008): Regression discontinuity Designs: A Guide to Practice, Journal of Econometrics, 142(2), 615635.

[68] Kramon, Eric and Daniel N. Posner. 2013. Who Benefits from Distributive Politics? How the Outcomes One Studies Affect the Answers One Gets. Perspectives on Politics 11(2).

[69] Keulder, C. (2000): Traditional Leaders, State, Society and Democracy: A Reader in Namibian Politics, Christiaan Keulder (Ed.) Macmillan Education Namibia

[70] Kum, Joseph M. (1993): The Central African Subregion, in Disarmament: Workshop on the Role of Border Problems in African Peace and Security (pp. 49-71). New York: United Nations

[71] La Porta, R. L., F. L. De Silanes, A. Shleifer and R. Vishny (1997): Legal Determinants of External Finance, Journal of Finance, 52(3), 11311150.

[72] La Porta, R. L., F. L. De Silanes, A. Shleifer and R. Vishny (1998): Law and Finance, Journal of Political Economy, 106(6), 11131155.

[73] Lange, M.K. (2004) British Colonial Legacies and Political Development. World Development, 32, 905-922.

[74] Long, J. D. (2012): The Determinants of Ethnic Voting, mimeo, UCL

[75] Lee, D. S., and T. Lemieux (2010): Regression Discontinuity Designs in Economics, Journal of Economic Literature, 48(2), 281355.

[76] Mamdani, M. (1996): Citizen and Subject. James Currey, London, UK.

[77] Mattes, Robert B. and Amanda Gouws (1999): Race, Ethnicity and Voting Behavior: Lessons from South Africa, In Electoral Systems and Conflict in Divided Societies, Andrew Reynolds and Ben Reilly (Eds.) Washington, DC: National Academy Press.

[78] Mattes, Robert B. and Jessica Piombo (2001): Opposition parties and the voters in South Africas General Election of 1999, Democratization 8(3): 101-12

[79] Mellinger, Andrew D., Jeffrey D. Sachs, and John L. Gallup (2000): Climate, Coastal Proximity, and Development. In The Oxford Handbook of Economic Geography, edited by Gordon L. Clark, Maryann P. Feldman, and Meric S. Gertler, 16994. New York: Oxford University Press

[80] Miguel, Edward (2004): Tribe or Nation? Nation-Building and Public Goods in Kenya versus Tanzania World Politics, Volume 56, Number 3, pp. 327-36

[81] Michalopoulos, S.(2012): The Origins of Ethnolinguistic Diversity, American Economic Review 81(1): 1508-39

[82] Michalopoulos, S. and E. Papaioannou (2011a): Divide and Rule or the Rule of the Divided? Evidence from Africa, NBER Working Paper 17184.

[83] Michalopoulos, S. and E. Papaioannou (2011b): The Long-Run Effects of the Scramble for Africa, NBER Working Paper 17620

[84] Michalopoulos, S. and E. Papaioannou (2012): National Institutions and African Development: Evidence from Partitioned Ethnicities, NBER Working Paper 18275.

[85] Michalopoulos, S. and E. Papaioannou (2013): Pre-colonial Ethnic Institutions and Contemporary African Development Econometrica, 2013, 81(1): 113152.

[86] Michalopoulos, S. and E. Papaioannou (2014a): National Institutions and Subnational Development in Africa Quarterly Journal of Economics, 2014, 129 (1): 151-213

[87] Michalopoulos, S. and E. Papaioannou (2014b):"On the Ethnic Origins of African Development Chiefs and Pre-colonial Political Centralization" NBER Working Paper No. 20513

[88] Miguel, Edward and Farhan Zaidi (2003) "Do Politicians Reward their Supporters? Regression Discontinuity Evidence from Ghana." Unpublished manuscript, University of California Berkeley.

[89] Montalvo, Jos G. and Marta Reynal-Querol (2005). Ethnic Polarization, Potential Conflict and Civil Wars American Economic Review, 95(3): 796816

[90] Moser, Christine. (2008): "Poverty Reduction, Patronage or Vote Buying? The Allocation of Public Goods and the 2001 Election in Madagascar." Economic Development and Cultural Change 57 (1): 137-162

[91] Munshi K. (2014) "Community Networks and the Process of Development," Journal of Economic Perspectives, 28: 4976

[92] Murdock, G. P. (1959): Africa: Its Peoples and Their Culture History. McGraw-Hill Book Company, New York.

[93] Murdock, G. P. (1967): Ethnographic Atlas. University of Pittsburgh Press, Pittsburgh, PA

[94] Naidu, S., (2010): Suffrage, Schooling and Sorting in the Post-Bellum U.S. South, mimeo, Harvard University

[95] Norris, Pippa and Robert Mattes (2003). Does Ethnicity Determine Support for the Governing Party? Afrobarometer Paper no. 26

[96] North, Douglass C. (1990): Institutions, Institutional Change and Economic Performance. New York: Cambridge University Press

[97] Nugent, Paul (1996): Arbitrary Lines and the Peoples Minds: A Dissenting View on Colonial Boundaries in West Africa, in Paul Nugent A. I. Asiwaju (Eds.), African Boundaries (pp. 35-67). London: Pinter

[98] Nunn, N. (2008): The Long Term Effects of Africas Slave Trades, Quarterly Journal of Economics, 123(1), 139176.

[99] Nunn, N. and L. Wantchekon (2011): The Slave Trade and the Origins of Mistrust in Africa, American Economic Review, 101(7), 32213252.

[100] Odugbemi, Sina. (1995, April 3). Consensus and Stability. West Africa, 501-503.

[101] Ottaway, Marina. (1999, February 25). Keep out of Africa. Financial Times. Retrieved March 8, 1999, from http://www.ft.com/search9/cgi/vtopic

[102] Posner, Daniel (2004a), The Political Salience of Cultural Different: Why Chewas and Tumbukas are Allies in Zambia and Adversaries in Malawi, American Political Science Review (2004), 98: 529-545

[103] Posner, Daniel (2004b), Measuring Ethnic Fractionalization in Africa, American Journal of Political Science, 48: 849863

[104] Posner, Daniel. (2005): Institutions and Ethnic Politics in Africa. New York: Cambridge University Press

[105] Rainer, Ilia and Francesco Trebbi (2014): New Tools for the Analysis of Political Power in Africa, in African Successes: Government and Institutions, edited by Sebastian Edwards, Simon Johnson, and David N. Weil; University of Chicago Press

[106] Rodrik, Dani, Subramanian, Arvind Trebbi, Francesco (2004): Institutions Rule: The Primacy of Institutions Over Geography and Integration in Economic Development, Journal of Economics Growth 2(6): 131-165

[107] Spolaore, E. and R. Wacziarg (2009): The Diffusion of Development, Quarterly Journal of Economics, 124(2), 469529.

[108] Touval, Saadia (1969): The Sources of Status Quo and Irredentist Policies, in Carl G. Widstrand (Ed.), African Boundary Problems (pp. 101-118). Uppsala, Sweden: Scandinavian Institute of African Studies

[109] White, Douglas (1987): Cross-Cultural Instructional Atlas Cultural Diversity Data Base; World Cultures Journal Volume 3-2

Figures

Figure 1: Ethnic and Partitioned Groups

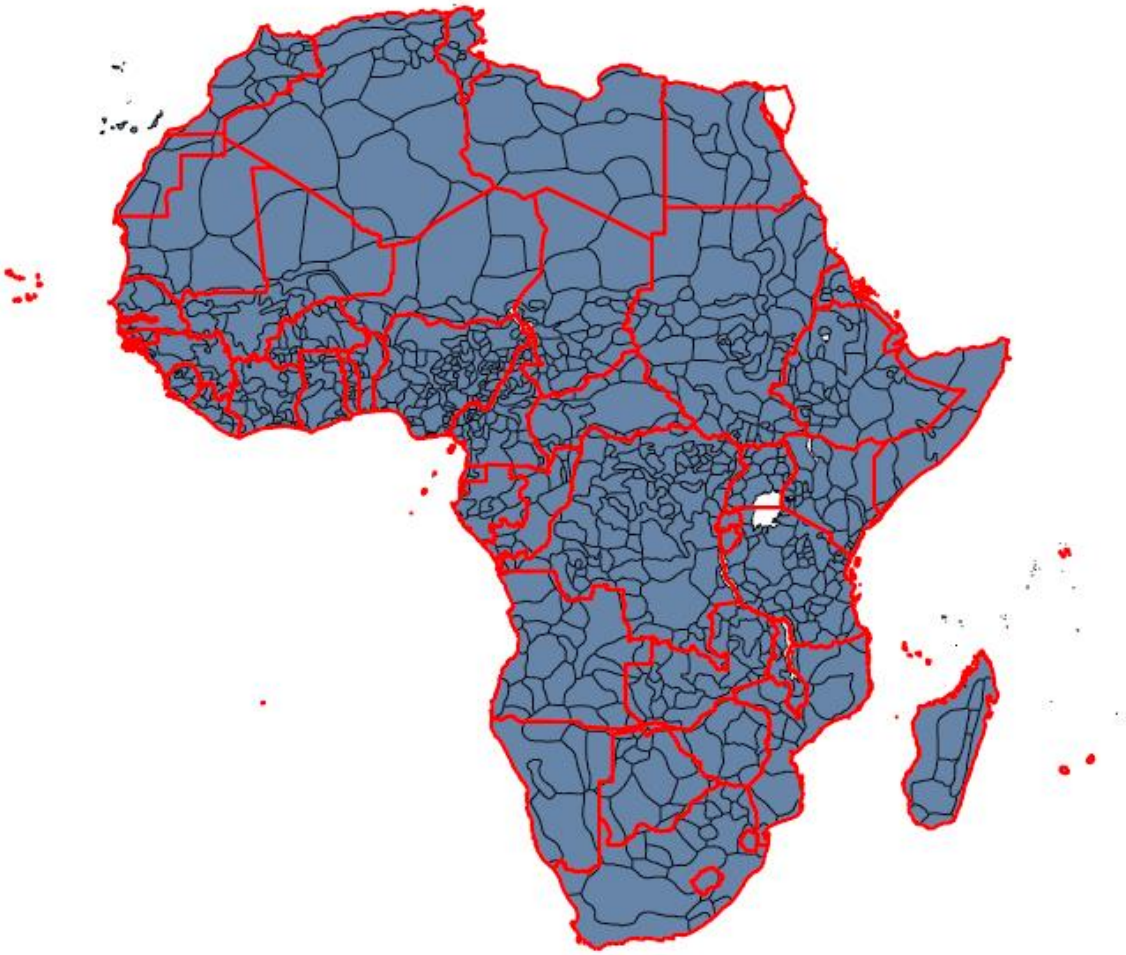


Figure 2: Partition of the Aulliminden Group

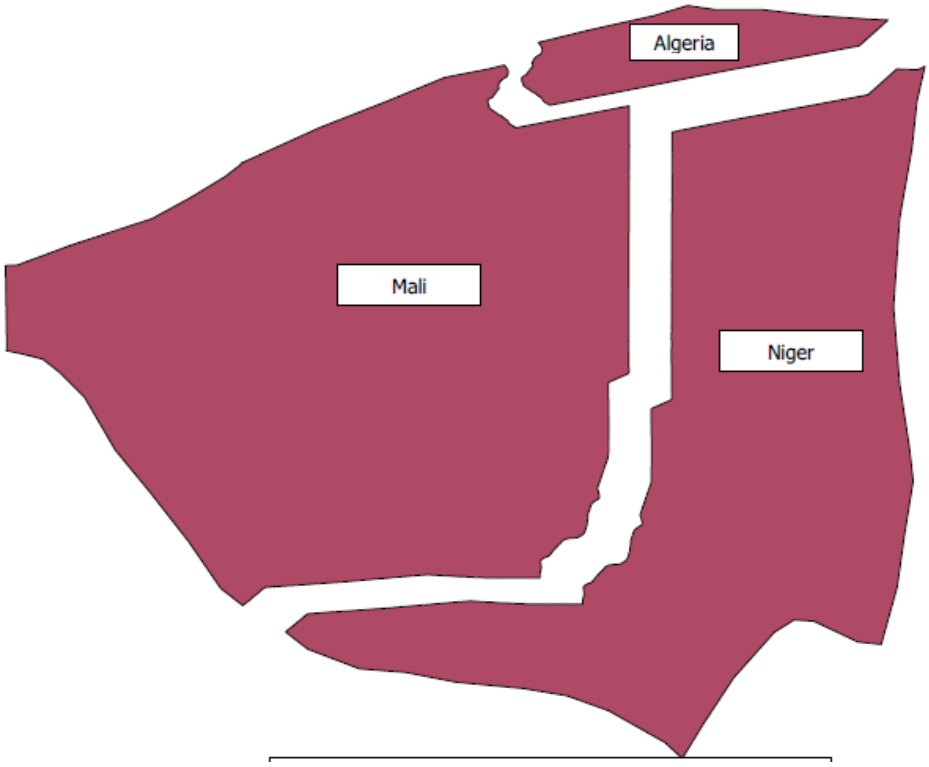


Figure 3: 1 km^2 Data on Total Economic Activity (converted into a Shapefile)

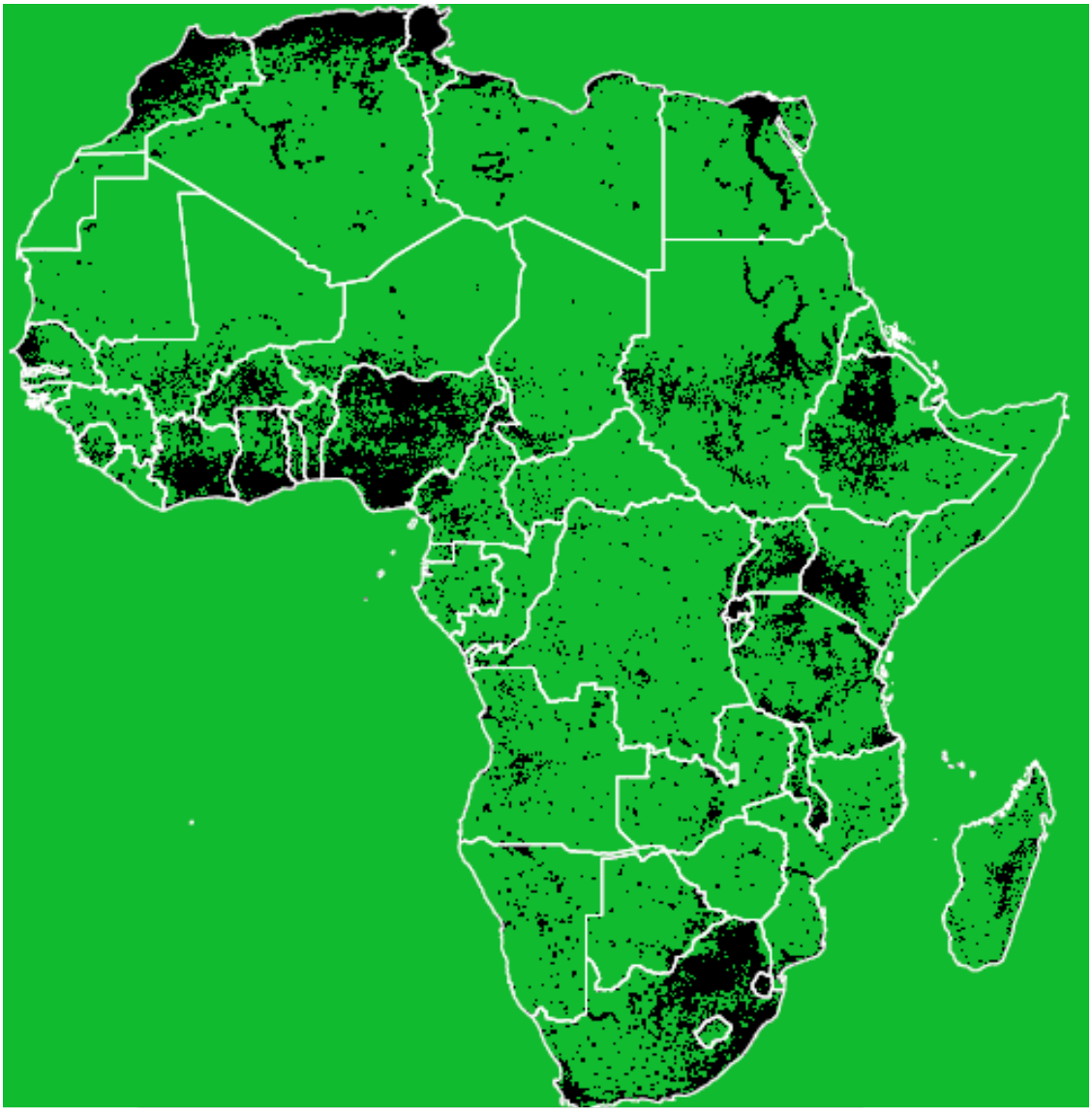
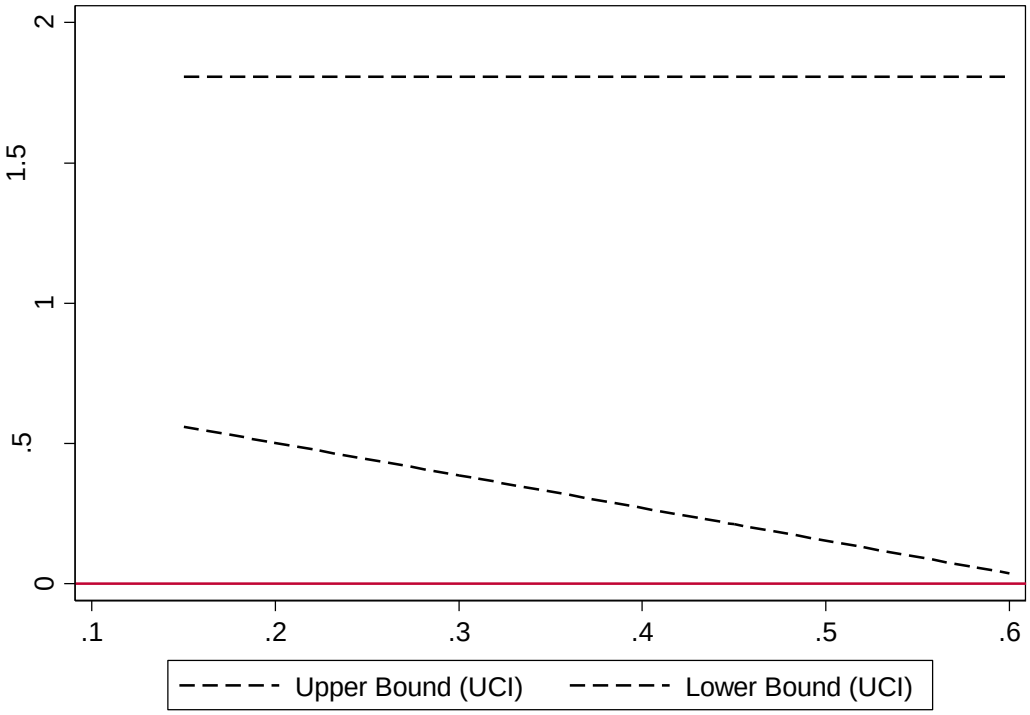


Figure 4: **Sensitivity to Violations of the Exclusion Restriction**



Tables

Table 1: Pairwise Correlation between Economic Activity and Light Density

	Mean Economic Activity	*Mean Light Density*
Mean Economic Activity	1.0000	
Mean Light Density (MP, 2013)	0.8899	1.0000

Table 2: Descriptive Statistics

Variable	*Obs*	*Mean*	*Std. Dev.*	*Min*	*Max*
Mean Economic Activity	1300	.0793029	.3057063	0	7.251074
Partition-Dummy	1300	.6384615	.4806309	0	1
Population Share Groups	1297	.0374967	.095127	0	.9865889
Size Partition Group/Country Size	1300	.0377034	.0813861	1.22e-08	.9343922
Mean Population Partitioned Group	1297	612611.1	2162498	0	6.15e+07
Mean Population Ethnic Homeland	1300	1018439	2424008	410.8929	6.15e+07
Size Part. Group/Pop. Ethnic Home.	1297	.6422514	.4031694	0	1
Size Part. Group/Ethnic Homeland	1300	.6407692	.3916778	1.71e-06	1
Population Density	1299	50.91497	93.00702	.0259535	1840.406

Table 3: Cross-Group Analysis

Dependent Var: Log (0.01+Mean Economic Activity)			
Estimation Method: OLS	Model 1	Model 2	Model 3
Partitioned Group Dummy	-0.112***	-0.0552	-0.0776
	(0.0390)	(0.0426)	(0.0626)
Group Population Share	1.703***	1.648***	1.600***
	(0.266)	(0.276)	(0.380)
Country Fixed Effects	YES	YES	YES
Observations	1,291	1,286	710
R-squared	0.744	0.756	0.790

Model 1 only includes country fixed effects.

Model 2 includes country fixed effects, population density in 2000, distance from the capital, mean slope, mountainous terrain, rugged terrain, distance from the sea, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability to crops

Model 3 include all the above plus dummies for settlement, for gathering dependence, and jurisdictional Hierarchy from Murdock (1959).

Double Clustered Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 4: **Group Share and Pre-Colonial/Geographical Characteristics**

Dependent Variables:	***Pop 1800***	***Malaria***	***Soil Suit.***	***Mountains***	***Water***	***Oil Fields***
Estimation Method: ***OLS***	***Model 1***	***Model 2***	***Model 3***	***Model 4***	***Model 5***	***Model 6***
Group Population Share	-1.029	0.0377	-0.139	-0.0490	-0.0298	-0.438
	(0.961)	(0.0497)	(0.0930)	(0.0389)	(0.0338)	(0.435)
Observations	827	827	827	827	827	827
R-squared	0.51	0.87	0.44	0.42	0.55	0.39

Dependent Variables are: Population density in 1800 (Model 1), Malaria Suitability (Model 2), Soil Suitability (Model 3), Mountainous Terrain (Model 4), Presence of Water Bodies (Model 5), and Onshore Oil Fields (Model 6).
The sample only includes partitioned groups. Ethnic fixed effects included
Double Clustered Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 5: Matched-Groups

Dependent Var: Log (0.01+Mean Economic Activity)				
Estimation Method:	***OLS***	***OLS***	***Spatial HAC***	**Spatial HAC**
	Model 1	**Model 2**	**Model 3**	**Model 4**
Group Population Share	1.451***	1.620***	1.620***	1.620***
	(0.393)	(0.455)	(0.283)	(0.324)
Ethnic Fixed Effects	YES	YES	YES	YES
Country Fixed Effects	NO	YES	YES	YES
Observations	816	816	816	816
R-squared	0.824	0.871	0.991	0.991

Additional controls include: Population density in 2000, distance from the capital, mean slope, mountainous terrain, rugged terrain, distance from the sea, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability to crops.

Double Clustered Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 6: Controlling for Surface Area, Total Population, and Alternative Functional form on GDP

Estimation Method: OLS	**Model 1**	**Model 2**	**Model 3**
Dependent Vars:	**Log (0.01+GDP)**	**Log (0.01+GDP)**	**Log (GDP)**
Group Population Share	1.196***	1.106**	2.037***
	(0.460)	(0.454)	(0.688)
Ethnic Fixed Effects	YES	YES	YES
Country Fixed Effects	YES	YES	YES
Observations	826	816	816
R-squared	0.879	0.883	0.858

Additional controls include: Population density in 2000, distance from the capital, mean slope, mountainous terrain, rugged terrain, distance from the sea, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability to crops. Model 1 also includes the log of the total surface area and Model 2 includes all the previous controls plus total population within the group. The latter controls are not included in Model 3.
Double Clustered Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table 7: Dropping Small Groups

Dependent Var: Log (0.01+Mean Economic Activity)		
Estimation Method: Spatial HAC	***Model 1***	***Model 2***
Partitioned Group Pop. Share	1.547***	2.188***
	(0.368)	(0.380)
Tribe Fixed Effects	YES	YES
Country Fixed Effects	YES	YES
Observations	571	483
R-squared	0.993	0.994
Sample	***Groups 5% of homeland***	***Groups 10% of homeland***

Additional controls include: Population density in 2000, distance from the capital, mean slope, mountainous terrain, rugged terrain, distance from the sea, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability to crops.
Conley (2008) HAC Spatial Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.
The distance threshold for both models is 100km.

Table 8: Regression Discontinuity

Dependent Var: Log (0.01+Mean Economic Activity)			
Estimation Method: OLS	***Model 1***	***Model 2***	***Model 3***
Group Population Share	1.556***	1.656***	1.434
	(0.478)	(0.531)	(0.999)
Ethnic Fixed Effects	YES	YES	YES
Country Fixed Effects	YES	YES	YES
Observations	816	728	563
R-squared	0.744	0.756	0.790
Method	Cubic Polynomial	Within 100km	Within 50km

Additional Controls Include: population density in 2000, distance from the capital, mean slope, mountainous terrain, distance from the sea, rugged terrain, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability.
In Model 1 we use a cubic polynomial on the distance from centroid distance of the group from the border.
In Model 2 and Model 3 we estimate the local average effect using a local linear regression discontinuity approach. In this case the distance from the border is set to 100km in Model 2 and 50km in Model 3.
Double Clustered Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 9: IV Estimates

Dependent Var: Log (0.01+Economic Activity)			
Estimation Method: IV	**Model 1**	**Model 2**	**Model 3**
Partitioned Group Pop. Share	1.272***	1.105*	1.831***
	(0.365)	(0.636)	(0.881)
	First-stage Stats and Diagnostic Tests		
Size Partitioned Group	0.8542***	0.8491***	0.6634***
	(0.0602)	(0.0786)	(0.1283)
Anderson Canon LR Stats	812.239	412.869	123.79
Cragg-Donal F-Stats	1371.448	592.951	137.600
Stock and Yogo Critical Value	16.38	16.38	16.38
Tribe Fixed Effects	YES	YES	YES
Country Fixed Effects	YES	YES	YES
Observations	816	571	483
R-squared	0.431	0.386	0.394
Sample	***Partitioned Groups***	***Groups 5% of homeland***	***Groups 10% of homeland***

Additional Controls Include: population density in 2000, distance from the capital, mean slope, mountainous terrain, distance from the sea, rugged terrain, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability

Double Clustered Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 10: Alternative Dependent Variables

Dependent Variable:	***Ever Been Without Cash***	***Ever Been Without Fuel***
Estimation Method: OLS	***Model 1***	***Model 2***
Partitioned Group Pop. Share	-0.940**	-0.720***
	(0.450)	(0.194)
Ethnic Fixed Effects	YES	YES
Country Fixed Effects	YES	YES
Observations	19,387	19,395
R-squared	0.27	0.13

Additional controls include:

1) Geographical variables: population density in 2000, distance from the capital, mean slope, mountainous terrain, rugged terrain, distance from the sea, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability to crops
2) District controls: whether in the PSU/EA there is electricity, piped water, a sewage system, cell phone service, post office, health clinic, police station, etc.
3) Individual controls: age, gender, occupation, education, support for democracy, country and personal economic conditions, trust others, etc.

Double Clustered Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 11: Partitioned Groups and Infrastructures

Dependent Variable:	***Distance from Road***	***Distance from Railroad***
Estimation Method: OLS	***Model 1***	***Model 2***
Partitioned Group Pop. Share	-4.699***	-7.291**
	(1.807)	(2.809)
Ethnic Fixed Effects	YES	YES
Country Fixed Effects	YES	YES
Observations	821	821
R-squared	0.68	0.91

Additional Controls Include: population density in 2000, distance from the capital, mean slope, mountainous terrain, distance from the sea, rugged terrain, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability

Double Clustered Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1

Table 12: Partitioned Groups and Informal Institutions

Dependent Variable: Ethnic Leaders Have Primary Resp. for:	***Collect Income tax***	***Solving Disputes***	***Maintain Rule and Law***	**Allocation of Land**
Estimation Method: OLS	**Model 1**	**Model 2**	**Model 3**	**Model 4**
Group Population Share	-0.0042	0.0676	-0.278**	-0.0923**
	(0.0296)	(0.104)	(0.125)	(0.0461)
Ethnic Fixed Effects	YES	YES	YES	YES
Country Fixed Effects	YES	YES	YES	YES
Observations	18,544	19,018	18,901	18,959
R-squared	0.04	0.21	0.20	0.07

Additional controls include:

1) Geographical variables: population density in 2000, distance from the capital, mean slope, mountainous terrain, rugged terrain, distance from the sea, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability to crops, group mean GDP.
2) District controls: whether in the PSU/EA there is electricity, piped water, a sewage system, cell phone service, post office, health clinic, police station, etc.
3) Individual controls: age, gender, occupation, education, support for democracy, country and personal economic conditions, trust others, etc.

Double Clustered Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table 13: Partitioned Groups and Trust

Dependent Variable:	***Trust Relatives***	***Trust Others***	***Trust Other Groups***
Estimation Method: OLS	**Model 1**	**Model 2**	**Model 3**
Group Share Population	-0.214	-0.488***	-1.229***
	(0.158)	(0.177)	(0.180)
Tribe Fixed Effects	YES	YES	YES
Country Fixed Effects	YES	YES	YES
Observations	19,416	19,373	19,258
R-squared	0.16	0.16	0.16

Additional controls include:

1) Geographical variables: population density in 2000, distance from the capital, mean slope, mountainous terrain, rugged terrain, distance from the sea, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability to crops, group mean GDP.
2) District controls: whether in the PSU/EA there is electricity, piped water, a sewage system, cell phone service, post office, health clinic, police station, etc.
3) Individual controls: age, gender, occupation, education, support for democracy, country and personal economic conditions, etc.

Double Clustered Standard errors in parentheses: *** p<0.01, ** p<0.05, * p<0.1.

Table Appendix

Table A1: Quadratic Form on Population Share

Dependent Var: Log (0.01+Mean Economic Activity)	
Estimation Method:	***OLS***
	Model 1
Group Population Share	4.447***
	(1.010)
Group Population Share Squared	-3.869***
	(1.391)
Ethnic Fixed Effects	Yes
Country Fixed Effects	Yes
Observations	816
R-squared	0.877

Additional controls include: Population density in 2000, distance from the capital, mean slope, mountainous terrain, rugged terrain, distance from the sea, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability to crops.
Double Clustered Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Table A2: Controlling for the Surface Area of the Group.

Dependent Var: Log (0.01+Mean Economic Activity)	
Estimation Method:	***OLS***
	Model 1
Group Population Share	2.213***
	(0.706)
Ratio Group Surface Area to Country Surface Area	-0.0804
	(0.753)
Ethnic Fixed Effects	Yes
Country Fixed Effects	Yes
Observations	816
R-squared	0.872

Additional controls include: Population density in 2000, distance from the capital, mean slope, mountainous terrain, rugged terrain, distance from the sea, onshore oil fields, population in 1800, malaria environment, water availability and soil suitability to crops.
Double Clustered Standard errors in parentheses *** p<0.01, ** p<0.05, * p<0.1

Data Description

Data Source		
Variables	Description	Source
Ethnic Groups Map		Murdock (1959,1967)
State Boundaries		Global Administrative Database (GADM)
Economic Activity	Satellite Imagery of light density from the National Geophysical Data Center (NOAA/NGDC)	Ghosh et al. (2010)
Population Density		Gridded Population of the World (GPW) - SEDAC
Partition Dummy	Intersection between state boundaries and ethnic groups location	Murdock (1959,1967) + GADM
Group Share	Group Area/Country Area	
Mountains /Terrains	Digital Elevation Model	FAO-GeoNetwork
Rugged Terrain	Digital Elevation Model	Computed using QGIS DEM model plugin
Water Availability	Water Basins	FAO-GeoNetwork
Distance from the Sea	Distance to the Nearest Coast	NASA Ocean Biology Processing Group
Distance from the Capital	Euclidean Distance from the Capital	CEPII (cepii.fr)
Population Density in 1800	History Database of the Global Environment	HYDE
Onshore Oil Fields	Number of Oil Fields/Group Surface Area	UCDP/PRIO
Environmental Suitability to Malaria	1km² Spatial Data from a biological model which incorporates the effect of climate on 1) vector lifespan and 2) the duration of *P. falciparum* sporogeny.	Oxford Atlas Malaria Project
Crop Suitability	Digital Soil Map	FAO GEONETWORK
Settlement Types, Dependence on Gathering and Juridical Hierarchy		Murdock (1959, 1967)
Roads and Railways	Spatial Data (Shape Lines)	FAO-GEONETWORK
Proxies for Informal Institutions	Questions Q58D, Q58E, Q58F, Q58H	The Afrobarometer (IV Round)
Proxies for Individual Income	Questions Q8D, Q8E	The Afrobarometer (IV Round)
Proxies for Trusts	Questions Q84A, Q84B, Q84C	The Afrobarometer (IV Round)
Regional Provision of Public Goods	Dummies for the provision of electricity, of piped water, a sewage system, health clinics, paved terrain, schools, whether there are any police or soldier stations, etc.	The Afrobarometer (IV Round)
Individual Effects	Dummies for age, sex, trust in democracy, employment status, trust in peers, education, urban, etc.	The Afrobarometer (IV Round)

Appendix II: Partitioned Ethnic Groups and Number of Countries

Name	Countries	Name	Countries
ABABDA	2	LAMBYA	3
ACHOLI	2	LENDU	2
ADAMAWA	3	LIGBI, DEGHA (SE)	2
ADARAWA	2	LIMBA	2
ADELE	2	LIPTAKO	2
AFAR	3	LOBI	2
AHAGGAREN	2	LOGO	2
ALGERIANS	2	LOMWE	2
ALUR	2	LOTUKO	2
AMBA	2	LUAPULA	2
AMBO	2	LUCHAZI	2
AMER	2	LUGBARA	3
AMHARA	2	LUMBO	2
ANA	2	LUNDA	2
ANUAK	2	LUNGU	2
ANYANG	2	LUO	3
ANYI	2	LUVALE	3
ARAD	2	MABA	2
ASBEN	2	MADI	2
ASSINI	2	MAKONDE	2
ATTA	2	MAKUA	2
ATYUTI	2	MALINKE	6
AULLIMINDEN	3	MAMBILA	2
AUSHI	2	MAMPRUSI	2
AVATIME	2	MANDARA	2
AZANDE	3	MANGA	2
AZJER	3	MANYIKA	2
BABUKUR	2	MASA	2
BAJUN	2	MASAI	2
BAKWE	2	MASALIT	2
BALANTE	2	MASHI	2
BAMBARA	2	MASINA	3
BANDA	3	MATAKAM	2
BANGI	2	MATENGO	2
BANYUN	2	MBAGANI	2
BANZIRI	2	MBERE	3
BARABRA	2	MBUKUSHU	3
BARARETTA	3	MBUNDA	2
BARGU	4	MEBAN	2
BASARI	2	MENDE	3
BASHI	3	MERARIT	2
BATA	2	MIJERTEIN	2
BAYA	2	MINIANKA	3
BERABER	2	MITTU	2
BERABISH	2	MOBA	4
BERIBERI	2	MOBER	2
BERTA	2	MOMBERA	2
BIAFADA	2	MOSSI	2
BIDEYAT	4	MPEZENI	2
BIRIFON	3	MUNDANG	2
BOBO	2	MUNDU	2
BOKI	2	MURLE	2
BONDJO	2	MUSGU	2
BONI	2	NAFANA	2
BORAN	2	NALU	2
BRONG	2	NAMA	2
BUDUMA	2	NARON	2
BUEM	2	NAUDEBA	2
BULOM	2	NDAU	2
BUSA	2	NDEBELE	2
BUSANSI	3	NDEMBU	3

BWAKA	3
CHAAMBA	2
CHAGA	2
CHAKOSSI	3
CHAMBA	2
CHEWA	3
CHIGA	3
CHOKWE	2
CHUABO	2
COMORIANS	2
DAFI	2
DAGARI	2
DAGOMBA	2
DAN	2
DARI	2
DAZA	2
DELIM	2
DENDI	3
DIALONKE	3
DIDINGA	3
DIGO	2
DIOLA	3
DOGON	2
DRAWA	2
DUI-MENIA	2
DUMA	2
DZEM	3
EGBA	3
EKOI	2
ESA	3
EWE	2
FAJULU	3
FANG	4
FIGIG	2
FILALA	2
FON	3
FOUTADJALON	4
FUNGON	2
FUR	2
GADAMES	3
GANDA	2
GERI	2
GIL	2
GISU	2
GOBU	2
GOLA	2
GOMANI	2
GREBO	2
GRUNSHI	2
GUDE	2
GUIN	2
GULA	2
GULE	2
GUMUZ	2
GUN	2
GURENSI	3
GURMA	4
GUSII	2
HA	2
HABBANIA	3
HADENDOWA	2
HAMAMA	2
HAMYAN	2
HAUSA	2
HAWIYA	2
HAYA	3
HEMAT	2
NDOGO	3
NDUKA	2
NEFUSA	2
NGALA	2
NGAMA	2
NGBANDI	2
NGERE	3
NGUMBA	2
NGWAKETSE	2
NGWATO	3
NKOLE	3
NSENGA	3
NSUNGLI	2
NUER	2
NUKWE	4
NUSAN	3
NYAKYUSA	2
NYANGIYA	2
NYANJA	2
NYASA	3
NYORO	2
NZANKARA	2
ODODOP	2
OGADEN	2
PANDE	2
PARE	2
POPO	2
PUKU	3
REGA	2
REGEIBAT	2
RENDILE	2
RESHIAT	3
RIYAH	3
ROLONG	2
RONGA	3
RUANDA	5
RUFFA	2
RUNDI	4
RUNGA	3
SAADI	2
SAB	2
SABEI	2
SAHO	2
SAMO	2
SANGA	3
SANUSI	2
SEGEJU	2
SEKE	2
SENUFO	3
SERER	2
SHAMBALA	2
SHASHI	2
SHEBELLE	2
SHILA	2
SHUWA	3
SIA	2
SILA	2
SINZA	2
SIWA	2
SOKOTO	2
SOMBA	2
SONGHAI	3
SONINKE	3
SONJO	2
SOTHO	2
SUBIA	4
SUNDI	2

HERERO	2
HIECHWARE	2
HLENGWE	3
HOLO	2
IBIBIO	2
IFORA	2
IMRAGEN	3
ISHAAK	2
IWA	2
JERID	2
JIE	2
KABRE	2
KAKA	2
KANEMBU	3
KANURI	2
KAONDE	2
KAPSIKI	2
KARA	2
KARAMOJONG	2
KARE	2
KEBU	2
KENTU	2
KGALAGADI	2
KGATLA	2
KHARGA	2
KISI	2
KISSI	3
KOBA	2
KOMA	2
KOMONO	2
KONGO	3
KONJO	2
KONKOMBA	2
KONO	2
KONYANKE	2
KORANKO	2
KOREKORE	3
KOTA	2
KOTOKO	2
KOTOPO	2
KOYAM	2
KPELLE	3
KRAN	2
KREISH	2
KUKU	2
KULANGO	3
KUNDA	3
KUNG	2
KUNTA	2
KUNYI	2
KWANGARE	2
LAKA (ADAMAWA	3
LALA	2
LAMBA	2
SURI	2
SUSU	3
SWAZI	3
TABWA	2
TAJAKANT	4
TAMA	2
TAWARA	2
TEDA	3
TEKE	3
TEKNA	2
TEM	2
TENDA	2
THONGA	3
TIENGA	3
TIGON	2
TIGRINYA	3
TIV	2
TLHARU	2
TLOKWA	3
TOMA	2
TONGA	2
TOPOTHA	3
TORO	2
TRIBU	2
TRIPOLITANIANS	2
TUBURI	2
TUKULOR	2
TUMBUKA	2
TUNISIANS	2
TURKANA	2
UDALAN	3
VAI	2
VENDA	2
VERE	2
VILI	4
WAKURA	2
WANGA	2
WIDEKUM	2
WOLOF	2
WUM	2
XAM	2
YAKA	2
YAKOMA	2
YALUNKA	2
YAO	3
YOMBE	3
ZAGHAWA	2
ZEKARA	2
ZENEGA	2
ZERMA	2
ZIMBA	2
ZULU	2
ZUMPER	2
Total	830